1979

A Discussion of an Empirical Bayes Multiple Comparison Technique

Donna Baranowski
Utah State University

A DISCUSSION OF AN EMPIRICAL BAYES MULTIPLE COMPARISON TECHNIQUE

by

Donna Baranowski

A report submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Plan B

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1979

Table of Contents

Page

List of Tables iii

List of Figures iv

Abstract v

Chapter

I. Introduction 1

II. Decision Theory 3

Bayesian Posterior Distribution 4

III. Estimation 6

The Bayes Estimator 7
The Empirical Bayes Estimation 10
An Application of Empirical Estimation of λ . 22

IV. Empirical Bayesian Multiple Comparison Procedures 26

Waller-Duncan k-ratio t-test 31
Example 33

V. Conclusion 36

References 38

Appendices 40

Appendix A. Magnesium Data 41
Appendix B. R Values 43
Appendix C. Q Values 45

Vita 47

List of Tables

Table		Page
1.	Nitrogen Values	8
2.	Quantities for calculating Empirical Bayes estimators .	23
3.	James-Stein Method	25
4.	Empirical Bayes critical t values, t_c	31
5.	Application of various multiple comparison techniques to the Black Gram data of Appendix A	35
6.	Minimum significant (α=.05) differences for the Black Gram data of Appendix A for the Student-Newman-Keuls (SNK) and Duncan's Shortest Significant Range (SSR) multiple comparisons.	35

List of Figures

Figure		Page
1.	Frequency distribution of nitrogen values	9
2.	Bayes estimates Lambda vs. $\overline{X}$	11
3.	Minimized expected Bayes Risk vs. amount of prior data	24
4.	Critical t values vs. prior odds ratio	32

Abstract

A Discussion of an Empirical Bayes

Multiple Comparison Technique

by

Donna Baranowski, Master of Science

Utah State University, 1979

Major Professor:
Department: Applied Statistics and Computer Science

This paper considers the application and comparison of Bayesian and nonBayesian multiple comparison techniques applied to sets of chemical analysis data. Suggestions are also made as to which methods should be used.

KEY WORDS: Bayesian Analysis; Multiple comparison; Estimation.

(54 pages)

Chapter I

Introduction

Statistics is an evolving science which sometimes lacks uniformity of language but tries to express truth and to make accurate predictions. The use of statistical techniques contributes to a better understanding of natural phenomenona by quantifying and bounding stochastic variables.

Researchers often choose comparison methods to analyze the significance of multiple treatment means. One that is used infrequently is Bayes method. This report is written to better inform the reader of the technique and significance of a Bayesian multiple comparison method.

The controversy between Bayesians and non-Bayesians does not stem from the principle of Bayes' criterion but from the prior probabilities that must be specified. The major points of controversy between Bayesian methods and classical statistical methods concern the choice of the prior density function and the necessity of additional assumptions. Most classical statisticians do not accept the notion of random parameters required for a Bayesian framework. When the parameters of a distribution are unknown, the Bayes estimators require

a prior distribution for these parameters. These parameters are then treated as random variables rather than constants.

In the following pages some examples of Bayesian estimation methods are derived and applied.

Chapter II

Decision Theory

A problem of choosing between two alternatives courses of action often arises in many disciplines. If all the facts are known, then the problem is lessened. Unfortunately, a decision is often required based on somewhat less than full information.

Statistical decision problems usually involve the use of data as an aid to decision making. The usual approach to treating decision problems involving data is to reduce their solution to the solving of a no-data problem. The problem of making a decision in the absence of data will first be considered followed by a problem with data.

In order to facilitate the following discussion, the notation and basic theory will now be given. The first things needed are the concepts of prior and posterior distributions.

Consider a random experiment having several events, say E_1 E_2, . . . , E_n; of which at most one event may occur. Also suppose that the judgmental probabilities of each event, $P[E_1]$, $P[E_2]$, . . . , $P[E_n]$ have been obtained. These are referred to as prior probabilities, because they represent the chance of the event occuring before the results from the empirical investigation are obtained. The investigation itself may have several possible outcomes or results, each of which

may be statistically dependent upon the events. A result is denoted as R and the conditional probabilities, $P[R|E]$, are often available. The result itself may be used to revise the probability of the events, since certain results may be more likely to follow certain events. These values are called posterior probabilities, since they revise the prior information and are calculated after obtaining the data.

An assignment of probability to events can be defined as a probability distribution. The total probability is distributed or assigned to the points and regions of the sample space, according to the relative likelihood of occurrence. The prior distribution is the function which represents the assignment of prior probabilities to points in the space before the data has been collected.

Bayesian Posterior Distribution

A basic belief held by Bayesian statisticians is that the state of nature can be described by the prior probability distribution. This prior distribution is the premise or foundation of Bayesian estimation. The posterior distribution represents the statistician's (present) interpretation about a particular distribution, given or conditional upon the observed data. Bayes' theorem modifies past or prior information by incorporating present information from a sample. The new information better reflects the distribution of the function of the random variable.

A random sample $x_1, \ldots, x_n$ from the density function X given or parameterized by λ, denoted $f(x|\lambda)$, will be used to estimate

the true value. Under the Bayesian framework, λ is a random variable with the distribution function $\pi(\lambda)$. The function of x given λ, $f(x|\lambda)$ is then a conditional probability density of X given λ. Note that the joint density function of X and λ is given by:

$$f(x, \lambda) = \pi(\lambda) f(x|\lambda). \qquad (2.1)$$

The posterior distribution is then defined as the conditional distribution of λ given x, $h(\lambda|x) = f(x, \lambda) / g(x)$, where $g(x)$ is defined as the marginal density of X. This conditional density function of λ given x is defined as the posters or probability distribution of λ given the sample.

The Bayes estimator of λ corresponding to the prior distribution function $G(\lambda)$ is the random variable $\phi_G(x)$ defined by the function:

$$\phi_G(x) = \frac{\int \lambda f(x\ \lambda)\, dG(\lambda)}{\int f(x\ \lambda)\, dG(\lambda)} \qquad (2.2)$$

which is the expected value of the posterior distribution of λ given $X = x$.

Chapter III

Estimation

In most cases, the more numerous the observations and the less variable the data, the closer the estimate of the parameter will approach the truth. Using random samples to estimate the value of a population parameter is one of the most common statistical methods. Estimation of a parameter requires that the user consider many different estimators before a particular procedure is chosen. A major concern in adopting a particular method of estimation is the accuracy and precision of the method. The estimator should not be subject to large variation and it should be close on average to the parameter it is estimating.

A numerically determined point estimate of a parameter λ can be viewed as a decision which can be correct or incorrect, depending on whether the estimate is actually equal to λ or not. Since the probability of an estimate being equal to the parameter is zero for continuous variables, a measure of the seriousness of the difference between the true value λ and the point estimate, $w(x)$ would be useful. The loss function will be used to quantify the severity of the consequences of taking a certain action. Assume that for each combination of state λ and action $\underline{a}$, there is a loss $L(\lambda,\underline{a})$ giving the negative measure of utility due to the consequences of taking action $\underline{a}$ when

nature is in state λ. λ is taken as the parameter while $\underline{a}$ is the estimate or test decision. This loss function, $L(\lambda, \underline{a})$, is assumed to be known, but sometimes a good decision can be made knowing only certain aspects of the loss function. Common loss functions include $|a - \lambda|$, $(a - \lambda)^2$, etc. For further examples of loss functions in decision making see the paper by Duncan (1975).

The loss value, $L(\lambda, \underline{a})$, is the calculated loss to the researcher if action $\underline{a}$ is taken when λ is the true state of nature. The decision function $\underline{a}$ could be replaced with the function $w(x)$ which is the Bayes decision function or estimator for the given prior distribution. This makes the loss function $L(\lambda, w(x))$. It is reasonable to try to choose $w(x)$ such that $L(\lambda, w(x))$ is minimized. This minimization is rarely possible under a Bayesian framework since λ is a random variable. This leads to the idea of risk which is the expected value of the loss function. The expectation is over values of x for a fixed value of λ.

Since the Bayesian views λ as a random variable, the next objective is to select that value of λ which minimizes the expected loss or the risk. The value of λ which minimizes this risk is then defined as the Bayes estimator of λ. An example is presented in the following pages to help clarify the foregoing.

The Bayes Estimator

The data in Table 1 are a collection of nitrogen values gathered from prior analyses which were made daily using a Coleman Nitrogen

Analyzer II. The data in Table 1 were collected as part of a Desert Biome project during 1976 at Utah State University, under the direction of James MacMahon (Principal Investigator). A plot of the frequency distribution with the percent nitrogen ranging from 9.9 to 10.4 (0.1 intervals) is given in Figure 1.

Table 1
Nitrogen Values

10.230	10.304
10.060	10.166
10.134	10.283
10.187	10.365
9.912	10.314
10.113	10.134
9.912	10.155
9.975	10.007
10.049	10.102
10.155	10.261
9.922	10.198

A subjective decision was made, based on prior experience with chemical analyses of this type, to try a normal distribution for the X parameter. The parameter λ was taken as being uniformly distributed between the values of 9.0 to 10.6. The chosen or assumed interval corresponds with the range of individual observed x values. Where

$$f(x|\lambda) = (2x\sigma^2)^{-\frac{1}{2}} \exp\left[-(x-\lambda)^2/(2\sigma^2)\right] \qquad (3.1)$$

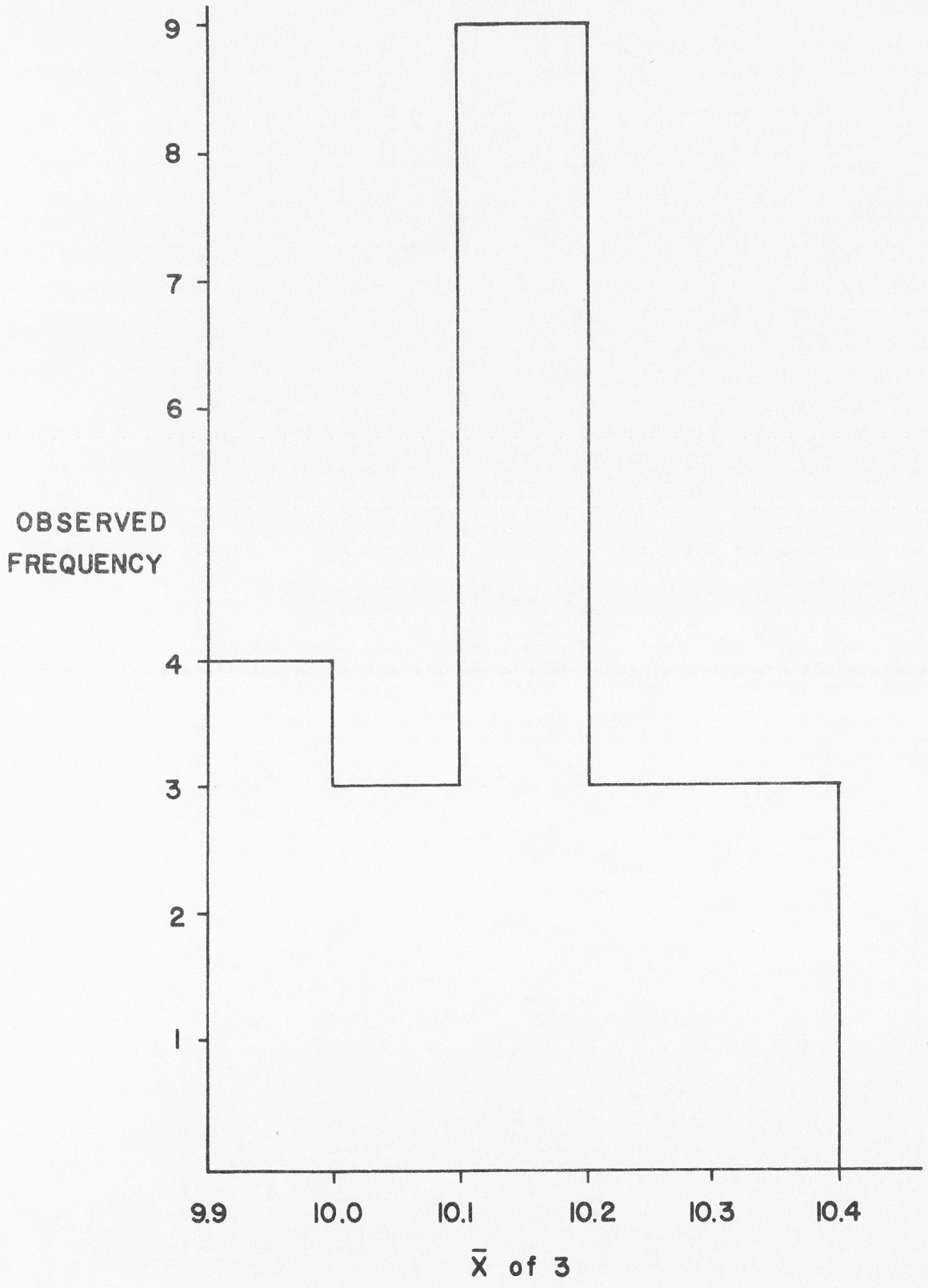

Figure 1. Frequency distribution of nitrogen values.

is substituted in equation (2.2). This leads to the equation:

$$\phi_G(x) = \frac{\int_9^{10.6} \lambda f(x|\lambda)\, d\lambda}{\int_9^{10.6} f(x|\lambda)\, d\lambda} \qquad (3.2)$$

A computer program was written to evaluate equation (3.2) using the trapezoidal rule for each $\bar{X}$. A plot of the estimates of λ, vs the $\bar{X}$'s is shown in Figure 2. The maximum value of λ over the range of data is shown in Figure 2 to be approximately 8.95. The estimates λ vary from approximately 8.95 to about 8.45 for $\bar{X}$'s between the values of 9.0 and 10.5 respectively.

Empirical Bayes Estimation

Since one of the main complaints of classical statisticians about Bayesian techniques is the need for more assumptions, a partial solution seems to be the concept of Empirical Bayes Estimation. The empirical adjective indicates that we will use past or historical data to estimate the prior distribution (or its parameters), and then use this estimated prior distribution in Bayesian methods.

As an example of Bayes estimation, the following estimator was derived in an unpublished paper by Lowe and Boes (1971). In this paper, the authors derive an empirical Bayes point estimator of the mean, assuming a normal prior distribution for λ of $N(\mu,\tau^2)$. Their first example gives the Bayes estimator of λ as:

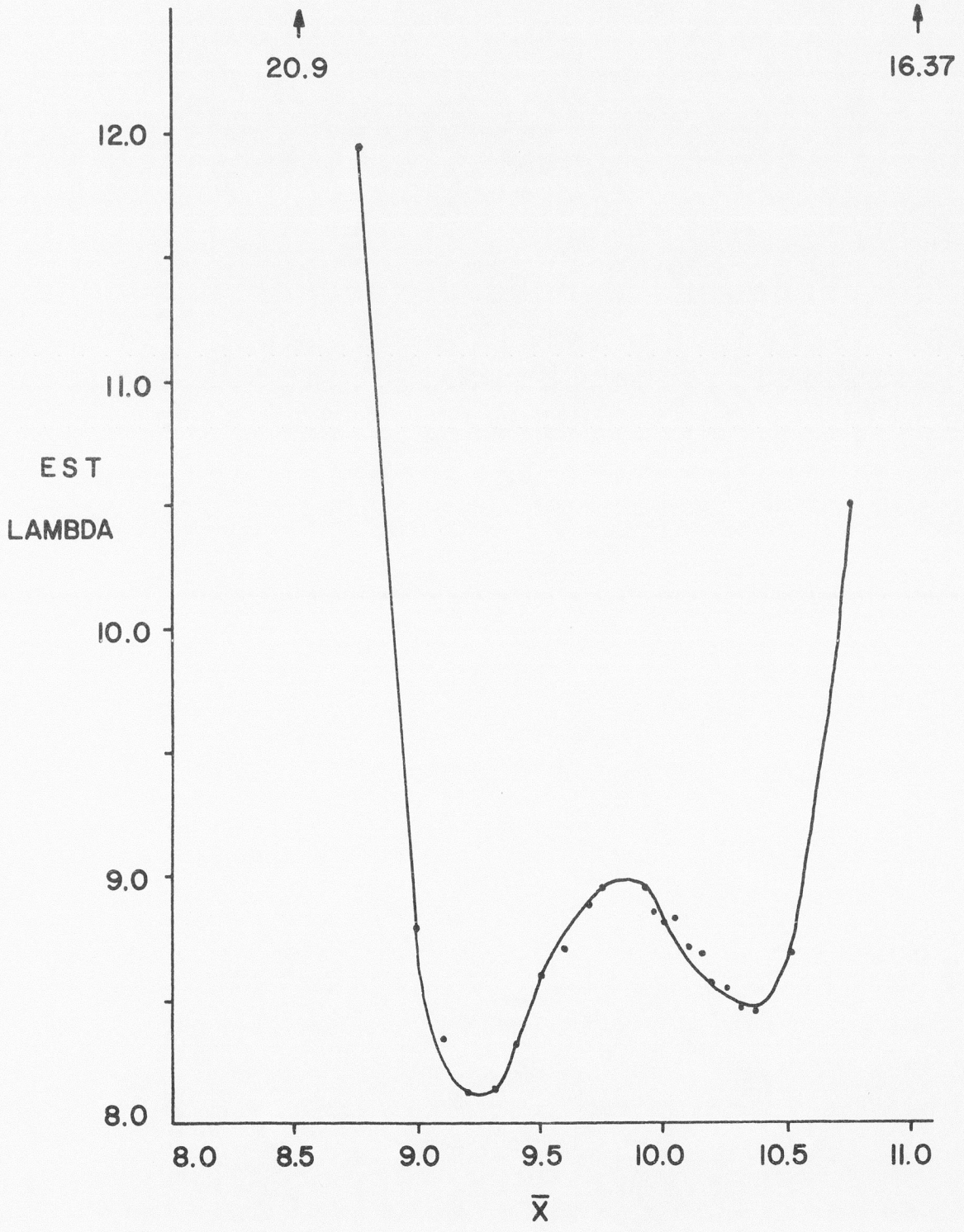

Figure 2. Bayes estimates Lambda vs. $\overline{X}$.

$$T_B = \frac{(n/\sigma^2)\bar{X} + (1/\tau^2)\mu}{(n/\sigma^2) + (1/\tau^2)} \qquad (3.3)$$

where σ^2, τ^2, and μ are assumed known. T_B is the mean of the prior distribution and is the estimator having the smallest risk with respect to a squared-error loss function, $L(\lambda, T_B) = (\lambda - T_B)^2$.

The risk function is based on the analysis of the expected value of the loss, where risk is the mean-squared-error and Bayes risk is the expected mean-squared-error (with respect to the prior distribution). From this the Bayes risk of the Bayes estimator of λ, T_B, is given as:

$$\frac{\sigma^2}{n} \quad \frac{n/\sigma^2}{n/\sigma^2 + 1/\tau^2} \qquad (3.4)$$

Because the empirical Bayes estimator is dependent on prior data with added assumptions, its Bayes risk is also dependent on the information. Therefore, the empirical Bayes estimator approaches optimization as the number of independent random samples increases and the empirical Bayes risk decreases.

It can be easily shown that the Bayes risk of T_B is smaller than the Bayes risk of $\bar{X}$, which is σ^2/n. Comparing the two Bayes risk formulas we see

$$\frac{\sigma^2}{n} > \frac{\sigma^2}{n} \quad \frac{(n/\sigma^2)}{(n/\sigma^2) + (1/\tau^2)} . \qquad (3.5)$$

Clearly there would be a greater risk using the estimator $\bar{X}$ than using the empirical Bayes estimator.

Consider the following example with the added assumption that the ratio of the sample and prior variance is known so that $\sigma_j^2/\tau^2 = \delta_j$. Substituting this into T_B gives

$$T_B = \left[\frac{n}{n+\delta_j}\right] \bar{X} + \left[\frac{\delta_j}{n+\delta_j}\right] \mu, \qquad (3.6)$$

where the only unknown parameter is μ. This suggests replacing μ by some estimate, say $\hat{\mu}$, based on past data. The following equation can then be used as an empirical Bayes estimator,

$$T_W = w\bar{X} + (1-w)\hat{\mu}, \qquad (3.7)$$

both $\hat{\mu}$ and w are selected to minimize the expected Bayes risk. The parameter w is a weight for the present sample ranging between 0 and 1. This makes (1-w) the weight for the past data. The mean-squared-error of T_W is

$$w^2 (\sigma^2/n) + (1-w)^2 (\lambda-\hat{\mu})^2, \qquad (3.8)$$

and the Bayes risk of T_W can be shown to be

$$w^2 (\sigma^2/n) + (1-w)^2 \{\tau^2 + (\mu-\hat{\mu})^2\}. \qquad (3.9)$$

The expected Bayes risk of T_W, with respect to w is

$$w^2(\sigma^2/n) + (1-w)^2\{\tau^2 + E[(\mu-\hat{\mu})^2]\} \qquad (3.10)$$

As in the previous example our goal is to show that the empirical Bayes estimator approach yields the least risk. This can be accomplished by selecting $\hat{\mu}$ and w so that the expected Bayes risk of T_W is minimal.

To minimize the risk, first choose $\hat{\mu}$ to minimize the mean square error for $\hat{\mu}$, $E[(\mu-\hat{\mu})]$, and then choose w. The linear combination of observations $\sum_{j=1}^{k} w_j \bar{X}_j = \hat{\mu}$ is an unbiased estimate of μ if $\Sigma w_j = 1$. The particular combination that is most efficient is the one which minimizes

$$\text{var}(\Sigma w_j \bar{X}_j) = \Sigma w_j^2 \text{ var}(\bar{X}_j) = \text{var}(\bar{X})\,\Sigma w_j^2 \qquad (3.11)$$

or the one that minimizes Σw_j^2, subject to $\Sigma w_j = 1$. It can be shown by using Lagrange multipliers that the set of w_j's which minimizes (3.11) is given by:

$$w_j^* = \frac{[(\sigma^2_j/n_j) + \tau^2]^{-1}}{\sum_{i=1}^{k}[\sigma^2_i/n_i) + \tau^2]^{-1}} = \frac{n_j/(n_j + \delta_j)}{\sum_{i=1}^{k}[n_i/(n_i + \delta_i)]}. \qquad (3.12)$$

By substituting w_j^* for w, $\hat{\mu}$ becomes $\hat{\mu}^* = \sum_{j=1}^{k} w_j \bar{X}_j$ and the expected Bayes risk of T_W is given by

$$E[(\hat{\mu}^*-\mu)^2] = \sum_{j=1}^{k}[n_j/(n_j\tau^2 + \sigma_j^2)]^{-1}. \qquad (3.13)$$

The next problem is to select a w to minimize the expected Bayes risk of T_W. A minimum w represented as w* is given as:

$$w^* = \frac{\tau^2 + \{\sum_{j=1}^{k} [n_j/(n_j \tau^2 + \sigma^2)]\}^{-1}}{(\sigma^2/n) + \tau^2 + \{\sum_{j=1}^{k} [n_j/(n_j \tau^2 + \sigma^2_j)]\}^{-1}} \quad (3.14)$$

Dividing through by τ^2 gives:

$$w^* = \frac{1 + \{\sum_{j=1}^{k} [n_j/(n_j + \delta_j)]\}^{-1}}{(\delta/n) + 1 + \{\sum_{j=1}^{k} [n_j/(n_j + \delta_j)]\}^{-1}} \quad (3.15)$$

where $\delta = \sigma^2/\tau^2$ from the present sample. Substituting w* into the Empirical Bayes estimator given previously yields:

$$T_w{}^* = w^* \bar{X} + (1-w^*)\hat{\mu}^*. \quad (3.16)$$

The expected Bayes risk of $T_W{}^*$ is then $(\sigma^2/n)w^*$. Note that if $w^* = 1$, then $T_W{}^*$ is equal to the risk of the sample mean. If $W^* < 1$, then the expected Bayes risk of $T_W{}^*$ is smaller than that of the sample mean, which means that $T_W{}^*$ is a better estimator than $\bar{X}$ if better is defined to be an estimator with smaller expected Bayes risk.

Another estimator similar to that developed by Lowe and Boes which also incorporates the Empirical Bayesian philosophy, is referred to as The James-Stein estimator. It uses observed averages to estimate unobservable quantities. This estimate sometimes contradicts

the traditional statistical theory that no other estimation rule is uniformly better than the observed average. Stein's paradox concerns the use of observed averages to estimate unobservable quantities.

The initial step in applying the James-Stein method is to determine the grand average, denoted by the symbol $\bar{\bar{y}}$. The essential process in The James-Stein method is the "shrinking" of all the individual averages toward this grand average. This shrinking factor is designated as c. It is determined by the observed averages and is given by the equation

$$c = 1 - \frac{(k-3)\sigma^2}{\sum_{j=1}^{k} (\bar{y}_j - \bar{\bar{y}})^2} \qquad (3.17)$$

Here k is the number of individual averages, σ^2 is the population variance and $\sum_{j=1}^{k} (\bar{y}_j - \bar{\bar{y}})^2$ is the sum of the squared deviations of the individual averages $\bar{y}_j$ from the grand average $\bar{\bar{y}}$. The estimator is found through the following equation:

$$\bar{Z} = \bar{\bar{y}} + c\,(\bar{y}_j - \bar{\bar{y}}). \qquad (3.18)$$

If the shrinking factor is unity or one, the $\bar{Z}$ is reduced to being equivalent to the individual average.

The risk function for the James-Stein estimator is

$$E[(\lambda - \bar{Z})^2] \qquad (3.19)$$

This risk is less than that for the sample averages no matter what the true values of the λ's happen to be. The reduction of risk can be substantial when the number of means is larger than five or six. The estimator does substantially better than the averages only if the true means lie near each other. It does at least marginally better no matter what the true means are.

The James-Stein estimator is similar to that of Bayes's equation. The James-Stein procedure has one important advantage over Bayes' method. The James-Stein method can be employed without knowledge of the prior distribution, and there is no need to assume the means being estimated are normally distributed. There is one drawback to the James-Stein method: it increases the risk function by an amount proportional to 3/k, where k is again the number of individual averages. The additional risk is negligible when k is greater than 15 or 20 and tolerable for k as small as 9.

As a second example of estimating λ, also taken from the Lowe and Boes report, assume the samples have the same variance and sample size. The variance is assumed unknown, but all the past samples have the same size, $n_j = n$ $(j = 1, \ldots, k)$, same unknown variance $\sigma_j^2 = \sigma^2$ $(j=1, \ldots, k)$, and λ is the value of a random variable which is distributed normally, $N(\mu, \tau^2)$ where both μ and τ^2 are unknown. The prior data has the distribution $N(\lambda_j, \sigma_j^2)$ where y_{ji} is a random sample from this data.

The form of the empirical Bayes estimator is:

$$T_{\mathcal{Z}} = \mathcal{Z}\bar{y} + (1-\mathcal{Z})\bar{\bar{y}}, \qquad (3.20)$$

where $\bar{y}$ without a subscript is the mean of the present data, $\bar{\bar{y}}$ is the grand average of the past data, and $\mathcal{Z}$ is a statistic independent of $\bar{\bar{y}}$, but depending on the past data alone. Again $\mathcal{Z}$ will be chosen, as was w, to make the expected Bayes risk of $T_{\mathcal{Z}}$ minimal. The mean squared error of $T_{\mathcal{Z}}$, with $\mathcal{Z}$ and $\bar{y}$ fixed, is $\mathcal{Z}^2(\sigma^2/n) + (1-\mathcal{Z})^2(\lambda-\bar{\bar{y}})$, and similarly the Bayes risk is

$$\mathcal{Z}^2(\sigma^2/n) + (1-\mathcal{Z})^2[\tau^2 + (\bar{y}-\mu)^2]. \qquad (3.21)$$

Assuming independence of $\mathcal{Z}$ and $\bar{\bar{y}}$, the expected Bayes risk is:

$$E[\mathcal{Z}^2](\sigma^2/n) + E[(1-\mathcal{Z})^2](\tau^2 + \mathrm{Var}[\bar{\bar{y}}]). \qquad (3.22)$$

As in the previous example, the objective is to select a weight $\mathcal{Z}$. However, $\mathcal{Z}$ is now a statistic which depends on the past data yet is independent of $\bar{\bar{y}}$, making the Bayes risk minimal. The statistics $\bar{\bar{y}}$, $\Sigma\Sigma(y_{ji}-\bar{y}_{j\cdot})^2$ and $\Sigma\,\bar{y}_j-\bar{\bar{y}})^2$ are mutually independent. As shown in Graybill (1976), (3.23) below is a chi-square random variable with $k(n-1)$ degrees of freedom, $\bar{\bar{y}} \sim N[\mu, (\sigma^2/kn) + (\tau^2/k)]$,

$$[\Sigma\Sigma(y_{ji} - \bar{y}_{j\cdot})^2]/\sigma^2. \qquad (3.23)$$

The above assumptions make the expected Bayes risk

$$E[\bar{Z}^2](\sigma^2/n) + E[(1-\bar{Z})^2][\tau^2 + (\sigma^2/kn) + (\tau^2/k)]. \quad (3.24)$$

Further, $\Sigma(\bar{y}_{j.} - \bar{\bar{y}})^2 / [\sigma^2/n + \tau^2]$ can be shown to be a chi-square random variable with (k-1) degrees of freedom. Since σ^2 and τ^2 are unknown, the equation:

$$1 - \bar{Z} = \frac{(\sigma^2/n)}{[(\sigma^2/n) + \tau^2][(k+1)/k]} \quad (3.25)$$

is not acceptable. The next step is to find a statistic Z which approximates the right hand side of the above equation. Since the group sum of squares, the error sum of squares from a one-way analysis of variance and $\bar{\bar{y}}$ are all mutually independent. The ratio

$$\frac{\Sigma\Sigma(y_{ji} - \bar{y}_{j.})^2 / [\sigma^2 k(n-1)]}{\Sigma(\bar{y}_{j.} - \bar{\bar{y}})^2 / [(\sigma^2/n) + \tau^2](k-1)} \quad (3.26)$$

can be shown to be F-distributed with k(n-1) and k-1 degrees of freedom, and the expectation of $\Sigma\Sigma(y_{ji} - \bar{y}_{j.})^2 / \Sigma(\bar{y}_{j} - \bar{\bar{y}})^2$ is proportional to $\sigma^2/(\sigma^2/n) + \tau^2$. The equation: $(1 - \bar{Z}) = b[\Sigma\Sigma(y_{ji} - \bar{y}_{j.})^2 / \Sigma(\bar{y}_{j} - \bar{\bar{y}})^2]$, can be minimized, giving the expected Bayes risk as a function of b, say

$$b^* = \frac{k(k-5)}{n(k+1)[k(n-1)+2]} \quad . \quad (3.27)$$

The existence of the variance of $\hat{Z}$ requires that k be greater than five; otherwise there would be no weight on the prior data in the calculation of $T_{\hat{Z}^*}$.

Substituting b* for b yields the weighting of $\hat{Z} = \hat{Z}^*$ equal to:

$$\hat{Z}^* = 1 - b^* \left[\Sigma\Sigma (y_{ji} - \bar{y}_{j.})^2 / \Sigma (\bar{y}_{j.} - \bar{\bar{y}})^2 \right], \quad (3.28)$$

and the empirical Bayes estimator equals to:

$$T_{\hat{Z}^*} = \hat{Z}^* \bar{y} + (1 - \hat{Z}^*) \bar{\bar{y}}. \quad (3.29)$$

The expected Bayes risk of $T_{\hat{Z}^*}$ is given by

$$\frac{\sigma^2}{n} \left[1 - \left(\frac{\sigma^2}{\sigma^2 + n\tau^2} \right) \left(\frac{(n-1)\, k^2\, (k-5)}{(k-3)\,(k+1)\,[k(n-1)+2]} \right) \right] \quad (3.30)$$

which is derived on the following page. This is less than the Bayes risk of $\bar{y}$, which is σ^2/n. As k increases, $T_{\hat{Z}^*}$ becomes asymptotically optimal.

To compare the empirical Bayes estimator, we need the expected Bayes risk of $T_{\hat{Z}^*}$.

When

$$\hat{Z}^* = 1 - b^* \left[\Sigma\Sigma (y_{ji} - \bar{y}_{j.})^2 / \Sigma (\bar{y}_{j.} - \bar{\bar{y}})^2 \right]$$

substituting b* gives

$$\not{Z}^* = \frac{1 - k(k-5)\left[\Sigma\Sigma(y_{ji} - \bar{y}_{j.})^2 / \Sigma(\bar{y}_{j.} - \bar{\bar{y}})^2\right]}{n(k-1)\left[k(n-1) + 2\right]} \qquad (3.31)$$

where the expectation of:

$$\Sigma\Sigma(y_{ji} - \bar{y})^2 / \Sigma(\bar{y}_{j.} - \bar{\bar{y}})^2 \qquad (3.32)$$

is proportional to

$$\sigma^2 / [(\sigma^2/n + \tau^2)]. \qquad (3.33)$$

Therefore $\not{Z}^*$ can be simplified by substitution to:

$$\not{Z}^* = 1 - \frac{k(k-5)\sigma^2}{n(k+1)\left[k(n-1) + 2\right]\left[\sigma^2/n + \tau^2\right]} \qquad (3.34)$$

Rearranging equation (3.34) yields,

$$\not{Z}^* = 1 - \left(\frac{\sigma^2}{\sigma^2 + n\tau^2}\right)\left(\frac{k(k-5)}{(k+1)\left[k(n-1) + 2\right]}\right) \qquad (3.35)$$

Taking $T_{Z^*} = Z^* \bar{y} + (1 - \not{Z}^*)\bar{\bar{y}}$ as an empirical Bayes estimator, the expected Bayes risk of $T_{\not{Z}^*}$ is given by

$$\frac{\sigma^2}{n} = 1 - \left(\frac{\sigma^2}{\sigma^2 + n\tau^2}\right)\left(\frac{k(n-1)\,k(k-5)}{(k-3)(k+1)\left[k(n-1) + 2\right]}\right) \qquad (3.36)$$

Again, the Bayes risk of $\bar{y}$, which is σ^2/n, is larger than the Bayes risk of $T_{\not{Z}^*}$.

An Application of
Empirical Estimation of λ

An application of the previous process for determining the empirical Bayes estimator λ is shown for the following data. These data, shown in Table 2, are derived from Table 1. Triplets of days were taken, and the adjoining columns give quantities needed for computing pertinent statistics.

The ratio of $\Sigma\Sigma (y_{ji} - \bar{y}_{j.})^2$ to $\Sigma (\bar{y}_{j.} - \bar{\bar{y}})^2$ can be seen stabilizing as the number of prior data means increases as shown in Figure 3. This shows that each additional mean will have diminished influence on the ratio, and hence a smaller influence on the estimator as the number of samples increases.

The James-Stein method was next applied to the data found in Table 2 yielding the following statistics, where the grand average, $\bar{\bar{y}}$ is equal to 10.16 and the variance is equal to .0248.

Here k is again the number of unknown means, and c and $\bar{Z}$ are defined in equation (3.17) and (3.18) respectively. As the number of past means increases, the value of c diminishes and the influence of the grand average, $\bar{\bar{y}}$, increases, which is parallel to the effect of increasing the number of prior data means on the empirical Bayes estimator, as mentioned before.

The James-Stein procedure has one important advantage over the Lowe and Boes method since the James-Stein estimator can be

Table 2

Quantities for calculating Empirical Bayes estimators

Type of data	y_{ji}'s	$\bar{y}_{j.}$	$\Sigma\Sigma(y_{ji}-\bar{y}_{j})^2$	b*	Cumulative SSy_j	$\sum_j n_j \bar{y}_{j.}^2 - \frac{y^2}{n}$	Z*	T_{Z^*}
Prior	10.20, 10.32, 10.16	10.23	0.0139					
	10.17, 10,11, 10.05	10.11	0.0072					
	9.91, 10.20, 10,17	10.09	0.0509					
	10.03, 10.32, 10.24	10.20	0.0449					
	10.01, 10.00, 10.01	10.01	0.001					
	9.73, 10.26, 10.11	10.03	0.1493		.2663	.3178		
Recent	9.96, 9.83, 9.61	9.80	0.0626	0.02048	.3289	.5667	.9829	9.4595
	10.29, 10.13, 9.95	10.12	0.0579	.036458	.3868	.3695	.9788	9.6920
	9.84, 10.00, 9.80	9.88	0.0224	.049383	.4092	.4764	.9483	8.8484
	10.35, 10.31, 9.97	10.21	0.0872	.06000	.4964	.5413	.9368	8.9294
	10.18, 10.37, 10.36	10.30	0.0229	.068871	.5193	.4863	.9184	8.6380
	10.37, 10.20, 10.34	10.30	0.0165	.076389	.5358	.4061	.8907	8.0714
	10.06, 10.26, 10.18	10.17	0.0203	.082840	.5561	.6226	.9210	8.5681
	10.22, 10.32, 10.29	10.28	0.0053	.088435	.5614	.8987	.9417	9.0912
	10.52, 10.34, 10.26	10.37	0.0355	.093333	.5969	.8652	.9326	8.9887
	10.35, 10.20, 10.40	10.32	0.0217	.097656	.6186	1.161923	.9460	9.2150
	10.24, 10.02, 10.13	10.13	0.0242	.101499	.6428	1.16250	.9420	8.9540
	10.18, 10.14, 10.14	10.15	0.0011	.104938	.6439	.9620	.9277	8.6823
	10.06, 9.93, 10.03	10.01	0.0093	.108033	.6532	1.2204	.9407	8.8145
	10.13, 10.08, 10.10	10.10	0.0013	.110833	.6545	1.0228	.9274	8.6304
	10.09, 10.24, 10.42	10.25	0.0545	.113379	.7090	.8465	.9031	8.2741
	10.24, 10.15, 10.21	10.20	0.0042	.115702	.7132	.8548	.9035	8.2366

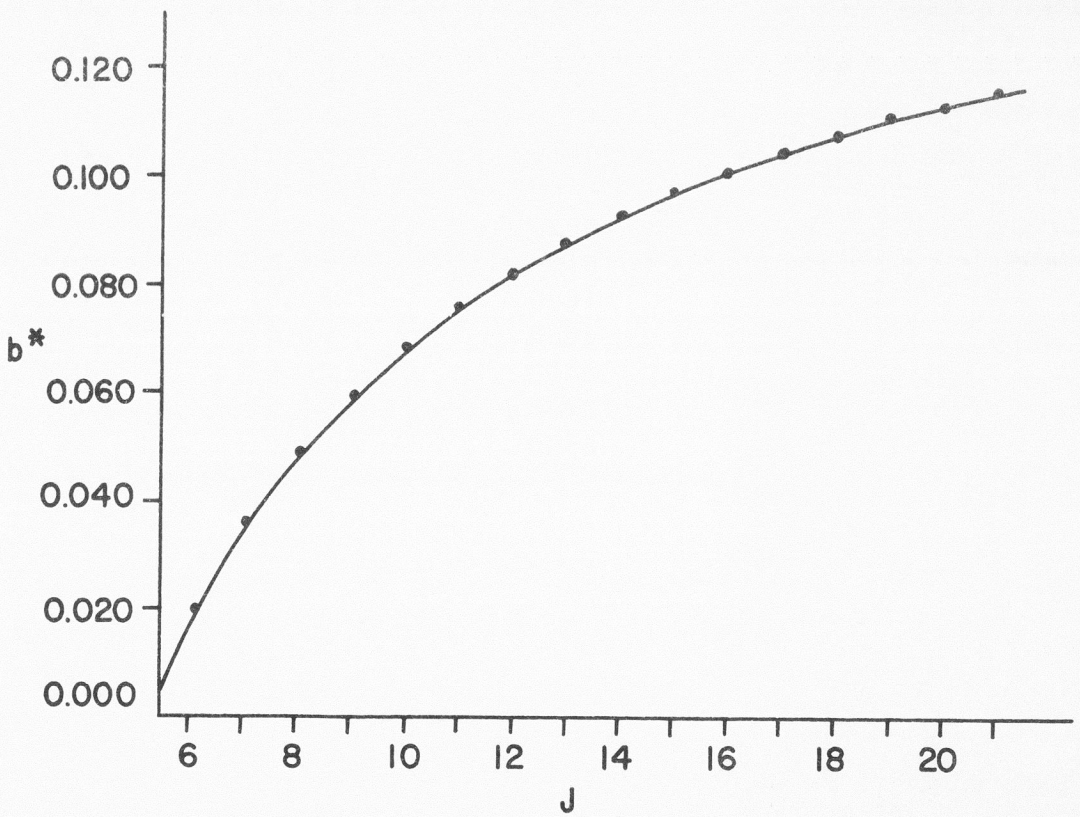

Figure 3. Minimized expected Bayes Risk vs. amount of prior data.

employed without knowledge of the prior distribution. In fact, one need not even assume the means being estimated are normally distributed.

Table 3
James-Stein Method

Triplet number (j)	$\bar{\bar{y}}$	$(\bar{y} - \bar{\bar{y}})^2$	k	c
7	10.11	.1310	1	--
8	10.16	.0018	2	--
9	10.12	.0794	3	--
10	10.17	.0023	4	.8833
11	10.18	.0191	5	.7858
12	10.18	.0191	6	.7032
13	10.16	.0001	7	.6045
14	10.18	.0140	8	.5317
15	10.19	.0433	9	.5170
16	10.18	.0250	10	.4788
17	10.16	.0010	11	.4460
18	10.16	.0001	12	.3784
19	10.14	.0231	13	.3096
20	10.15	.0038	14	.2485
21	10.17	.0078	15	.1974
22	10.17	.0014	16	.1339

Chapter IV

Empirical Bayesian Multiple Comparison Procedures

As a practical application of empirical Bayesian methodology, the following section deals with testing a comparison suggested by the data, first introduced by H. Robbins (1955).

Given n treatment means, $\bar{x}_1, \ldots, \bar{x}_n$ each based on r replications, a common problem is that of testing any comparison between means which may appear to be significant. The comparison may be a difference, $(\bar{x}_i - \bar{x}_j)$, between two of the means, which is a test of $H_0: \delta > 0$ against the one sided alternative $H_a: \delta > 0$, where δ is defined as the true difference between the population means, $\delta = \mu_1 - \mu_j$. This can easily be applied to a given set of prior equally plausible differences $d_1, \ldots, d_i$. to be tested for the hypothesis and the alternative $H_a: \delta_i > 0$ where the set is very large.

The incorrect decisions in choosing a hypothesis are referred to as Type I and Type II errors. A type I error is committed when a true H_0 is rejected. A Type II error is made if H_0 is accepted when it is actually false. It is conventional to denote the probabilities of these errors by α and β, respectively. If α is set equal to 5%, then the right-tailed 5% level t tests are applied to the differences simultaneously. Since the error rate is operative for each comparison, this is termed

a 5% level comparisonwise rule. A comparisonwise rule makes a Type I error in $100\alpha\%$ of the comparisons on the average. An experimentwise rule would allow a Type I error in only $100\alpha\%$ of all experiments on the average.

If a comparisonwise rule is devised which sets both α and β equal to 5%, two types of extreme results may occur. One extreme that is possible, Type A, results in only 5% of the tests being significant or only 5% of the null hypotheses being rejected. The other possible extreme, Type B, results in only 5% of the tests being not significant or only 5% of the null hypotheses not being rejected.

A dilemma encountered in multiple comparisons problems is that no approach consisting of simultaneous applications of several t tests can realistically hope to be acceptable when their error rate is specified based on a priori considerations alone. This approach is avoided because it reacts to the possibility of a Type A outcome by increasing the comparison t value or t_c, therefore decreasing the α for each prior test, although the actual outcome is Type A. If the outcome is intermediate, the increase in t_c should not have been made; if the outcome is Type B, the t_c should have been decreased.

In order to strike a compromise between Type A and Type B outcomes, an adequate rule for determining the significant t value must be allowed to depend on the overall outcome or significance for all the differences. The multiple comparison approaches of Fisher, Newman and Duncan achieve some of this dependence, but not nearly enough.

By never being less conservative than a comparisonwise rule, they too can fail to be as powerful as they should be for a Type B outcome. They can also fail to be sufficiently conservative for a Type A outcome.

A relatively new approach which recognizes and makes valuable use of these two simple identifying characteristics of multiple comparison problems is the additive losses concept developed by Duncan (1975). Additive losses are defined as the sum of the losses for the component decisions found in multiple comparison problems. Assuming there are n sample means, there are $n(n-1)/2$ pairwise comparisons to be made. The additive loss is then the sum of the losses for each of these possible decisions.

To illustrate the foregoing, consider the following problem originally presented by Duncan (1975). The number of differences, s, is equal to 1. Instead of choosing an α and β, a simple loss function is chosen. The loss function, $L(d_0|\delta)$, is defined as the loss when decision d_0 is taken but δ is the corresponding true difference between the means:

$$L(d_0|\delta) = \begin{cases} 0 & = \text{cost when } \delta = 0 \\ c_0 & = \text{cost when } \delta = \delta_a. \end{cases} \qquad (4.1)$$

Similarly, $L(d_a|\delta)$ is defined as

$$L(d_a \mid \delta) = \begin{cases} c_1 = \text{cost of rejecting } H_0 \text{ when } \delta = 0 \\ 0 = \text{cost of rejecting } H_0 \text{ when } \delta = \delta_a \end{cases} \quad (4.2)$$

The Bernoulli prior probability function is defined as the probability of δ taking on the values of 0 or δ_a;

$$P(\delta \mid p_a) = \begin{cases} p_a \text{ if } \delta = \delta_a \\ p_o \text{ if } \delta = 0 \end{cases} \quad (4.3)$$

where $p_0 = 1 - p_a$. Instead of seeking the most powerful α level test of H_0: $\delta = 0$ against H_a: $\delta = \delta_a$ with power $1 - \beta$, a Bayes rule or test is used which minimizes the Bayes risk:

$$B(r) = \sum_{i=0}^{a} \sum_{\delta=0}^{\delta_a} P(d_i \mid \delta)\, P(\delta \mid p_a)$$

$$= P(d_0 \mid \delta_a)\, c_0 p_a + p(d_a \mid 0)\, c_1 p_1. \quad (4.4)$$

Minimizing the above with Bayes rule by substituting the normal densities and taking the logarithms, it is found to be a right-tailed t value given by the equation:

$$t_c = \delta_a / 2 + (\ln c - \ln p)/\delta_a \quad (4.5)$$

where c is the loss or the seriousness ratio of a Type I to a Type II error, $c = c_1/c_0$ and p is the prior odds ratio in favor of the alternate hypothesis (H_a) or $p = p_a/(1 - p_a)$.

Therefore, if the costs of a Type I and Type II error are equal and the prior probabilities of H_a happening are equal to the probability of H_0 happening, then the t_c is equal to $\delta_a/2$.

An example of the empirical-Bayes additive losses approach when there is more than a single difference ($s > 1$), is given below. The additive losses result in the function:

$$c(d_i|\delta) = L(d_1|\delta_1) + \ldots + L(d_s|\delta_s). \qquad (4.5)$$

If the losses for a joint problem are the sums of the losses of its component problems, then the optimal rule for the sum is the same as that rule which minimizes loss for each of its components.

If the prior odds ratio p for H_a is unknown and if a difference, d, from a set of prior differences, p_a, approaches 0 (indicating a Type I result), then t_c will be very large and conservative. If p_a is near 1 (indicating a Type II result), t_c will be very small and powerful. Therefore, t is dependent on p_a. An example of this dependency is shown in Figure 4, derived from Table 4. As p nears zero or as p_a nears zero, t_c will be very large and conservative, indicating a Type I result. If p nears infinity, then p_a nears one and t_c will be very small and very powerful, or the likelihood of a Type II result is greater than that of a Type I error.

Table 4
Empirical Bayes critical t values, t_c.

p_a	p_0	$p = p_a/p_0$	t_c
0.1	0.9	0.111	2.4457
0.2	0.8	0.250	2.0397
0.3	0.7	0.430	1.7686
0.4	0.6	0.670	1.5468
0.5	0.5	1.000	1.3466
0.6	0.4	1.500	1.1439
0.7	0.3	2.330	0.9237
0.8	0.2	4.000	0.6534
0.9	0.1	9.000	0.2479

Waller-Duncan k-ratio t-test

A more widely applicable test for differences is the one by Waller and Duncan (1969). Critical values of the form $t_c = t_{(k, F, q, f)}$ are required where F is the F ratio for groups in the data set at hand, and f is the degrees of freedom for the between treatment mean square. The parameter k is the Type I to Type II error seriousness ratio, k_1/k_o and q is the amount of difference to be tested for. Since tables of $t_{(k, F, q, f)}$ are not available for arbitrary values of the parameters, the approximation

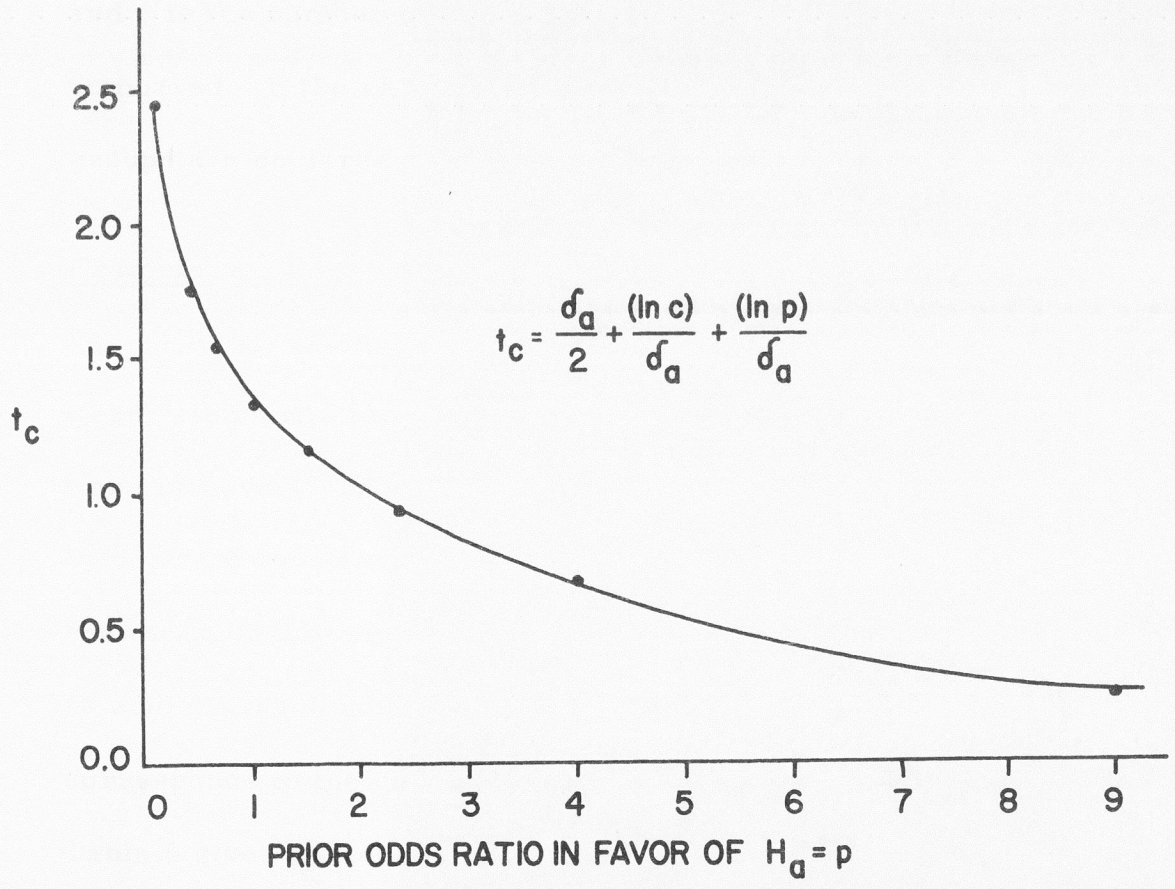

Figure 4. Critical t values vs. prior odds ratio.

$$t_c \approx (1 - 1/F)^{-\frac{1}{2}} \mathcal{Z}(k) \qquad (4.6)$$

is used where $\mathcal{Z}(k)$ is obtained from Table 6 of Duncan (1965). For two means to be declared significantly different using this procedure then requires $t > t_c$ where $t = (\bar{y}_i - \bar{y}_j)/s_d$ with

$$s_d = [2 s_e^2/r]^{\frac{1}{2}} \qquad (4.7)$$

where s_e^2 is the error mean square used:informing the sample F ratio and r is the number of observations in each of the two means being compared. If the calculated t exceeds the t_c given by (4.6), then the groups are declared statistically different.

Example

As an example of the foregoing technique, magnesium contents were determined for black gram several times after the seed coats were removed. The data along with their means and the analysis of variance are presented in Appendix A where the F ratio of 7.6103 is seen to be highly significant. Duncan (1965) suggests using k = 100 if no additional information on error seriousness is available since this corresponds to the usual $\alpha = .05$ type test. Using Duncan's (1965) Table 6 gives $t_c = (1 - 1/7.6103)^{-\frac{1}{2}}(1.721) = 1.84659$ to test say the difference between times 0 and 5, $t = (2.35167 - 2.38833/[2(.001128)/6]^{\frac{1}{2}} = 1.8904$. Since $t > 1.84659$, there is sufficient evidence to cause the rejection of $H_o: \mu_1 = \mu_2$.

This Waller-Duncan technique was applied to all pairwise differences by ranking the means and placing a continuous line beside homogeneous subsets in Table 5. Using this approach, two means must differ by at least .03581 to be significantly different. Several other multiple comparison techniques have also been included in Table 5 for comparison. The 5% LSD value for pairwise differences is .03919. Tukey's HSD value is .06349. Scheffé's value is .08043. Minimum difference values for the Student-Newman-Keuls (SNK) and Duncans shortest significant range (SSR) are presented in Table 6, and their differences are summarized in Table 5. From Table 5, it is clear that the Duncan-Waller technique is the most liberal, declaring 32 of the 36 pairwise comparisons significantly different. The LSD and the SSR are just slightly less liberal, declaring 31 of the 36 comparisons different. The SNK procedure is more conservative, declaring 29 of the 36 different. The HSD procedure, which uses the most conservative value of the SNK, declares only 27 significant differences while the Scheffé procedure detects only 24 differences.

In selecting one of these six procedures, one would almost certainly not suggest the use of either the HSD or the Scheffé approach since they are so conservative. While there is little difference between the LSD, the SNK and the SSR, the LSD (used after a significant F test) is the easiest of the three. The Waller-Duncan procedure is the most liberal of the set and is as easily applied as the LSD, HSD or Scheffé

procedure since only one critical value is needed. A good reference for all but the SNK procedure is Ott (1977).

Table 5
Application of various multiple comparison techniques to the Black Gram data of Appendix A.

Group	Ranked means	Duncan-Waller	LSD	SSR	SNK	HSD	Scheffé
3	2.095						
4	2.098						
7	2.140						
6	2.170						
5	2.232						
8	2.298						
1	2.352						
9	2.383						
2	2.388						

Table 6
Minimum significant (α=.05) differences for the Black Gram data of Appendix A for the Student-Newman-Keuls (SNK) and Duncan's Shortest Significant Range (SSR) multiple comparisons. Duncan-Waller (D-W), LSD and Scheffé's values are included for comparison

Number of Means	SNK	SSR	Others	
2	.0392	.0392	LSD	= .0392
3	.0472	.0413	D-W	= .0358
4	.0520	.0425	HSD	= .0635
5	.0554	.0435	Scheffé	= .0804
6	.0580	.0442		
7	.0602	.0448		
8	.0620	.0452		
9	.0635	.0457		

Chapter V

Conclusion

The controversy between Bayesians and non-Bayesians stems from the prior probabilities that are required of the Bayes method. It is important to realize that the use of Bayes techniques depends heavily on judgement and experience.

A basic belief of Bayes users is that the state of nature can be described by the probability distribution or prior distribution. The prior information is modified by incorporating present sample data which better represent the distribution function of the random variable.

The distribution function is used in the determination of an estimator. It was shown with this method that the Bayes risk of T_B is smaller than the Bayes risk of $\overline{X}$. Therefore, there would be a greater risk using the estimator $\overline{X}$ than using the empirical Bayes estimator.

A concern in using the Bayes method for estimation is the confidence the user has that it is the best method for the data. Consideration must also be given to the possibility of calculating an incorrect estimate.

The loss function is used for weighing the negative effect of taking a certain action because of an incorrect estimate. Data changes the problem of selecting an action to the selection of a decision function in view of a certain risk function.

Empirical Bayes estimation uses past data to estimate the prior distribution and then uses this prior estimate in Bayesian methods.

When the overall average of the decision errors is weighed, the Bayes risk accounts for every one of these errors, its loss factor, and its prior probability. This eliminates the need of choosing an appropriate α for each comparison example and analyzing the data in a non-Bayesian manner.

The Bayesian approach requires the user to specify a few more quantities, but in the end, the gain in control of comparisonwise and experimentive erros should far outweigh the added inconvenience of specifying more quantities.

References

Bernhardson, C. S. Type I Error Rates When Multiple Comparison Procedures Follow a Significant F test of ANOVA, Biometrics, 1975 31, 229-232.

Boardman, T. J. & Moffit. Graphical Monte Carlo Type I Error Rates for Multiple Comparison Procedures, Biometrics, 27, 738-745.

Carmer, S. G. & M. R. Swanson. Evaluation of Ten Pairwise Multiple Comparison Procedures by Monte Carlo Method, Journal of the American Statistical Association, 68, 66-74.

Cox, D. R. (May 1965), A Remark on Multiple Comparison Methods, Technometric, 7, 223.

Duncan, D. B. (June 1975), t Tests and Intervals for Comparisons Suggested by the Data, Biometrics, 31, 339-359.

Dunacn, D. B. (March 1955), Multiple-Range and Multiple-F Tests, Biometrics, 11, 1-42.

Duncan, D. B. (May 1965), A Bayesian Approach to Multiple Comparisons, Technometrics, 7:2, 171-222.

Grayhill, F. A. Theory and Application of The Linear Model, North Scituate, Mass:Duxbury Press, 1976.

Lindgren, B. W. Statistical Theory. 3rd ed. New York: Macmillan Publishing Co., Inc., 1976.

Lowe, V. W., Jr. & D. C. Boes. Empirical Bayes Point Estimation of the Mean of a Normal Distribution Assuming a Normal Prior. Unpublished ms. Colorado State University.

Mosteller, F. & D. L. Wallace. (1963), Inference in an Authorship Problem, American Statistical Association Journal, 58, 228-309.

O'Neill, R. & G. B. Wetherill. (1971). The Present State of Multiple Comparison Methods, Journal of the Royal Statistical Society Series B, 33, 218-233.

Ott, Lyman (1977), An Introduction to Statistical Methods and Data Analysis, N. Scituate, Mass: Duxbury Press.

Robbins, H. (1955), An Empirical Bayes Approach to Statistics, Proceedings of the 3rd Berkeley Symposium on Mathematics, Statistics and Probability, 1, 157-164.

Rutherford, J. R. & R. G. Krutchkoff, (1967), The Empirical Bayes Approach: Estimating the prior distribution, Biometrika, 54, 326-328.

Sprent, P. (March, 1970), Multiple Comparison Tests, Biometrics, 139-141.

Thomas, D. A. H. (1974), Error Rates in Multiple Comparisons Among Means--Results of a Simulation Exercise, Applied Statistics, 23, 284-293.

Waller, R. A. & D. B. Duncan. (December 1969), A Bayes Rule For the Symmetric Multiple Comparisons Problem, Journal of the American Statistical Association, 64, 1484-1503.

APPENDICES

Appendix A

Magnesium Data

Time (minutes)	Magnesium mg/g		Time means	Variances
0	2.38	2.29	2.35167	.00854
	2.44	2.24		
	2.47	2.29		
5	2.52	2.27	2.38833	.02554
	2.54	2.22		
	2.54	2.24		
10	2.17	2.02	2.09500	.00675
	2.17	2.02		
	2.17	2.02		
15	2.11	2.01	2.09833	.01530
	2.22	1.97		
	2.27	2.01		
20	2.21	2.31	2.23167	.00202
	2.21	2.21		
	2.19	2.26		
25	2.18	2.17	2.17000	.00012
	2.15	2.17		
	2.18	2.17		
30	2.08	2.19	2.14000	.00264
	2.08	2.19		
	2.13	2.17		
35	2.24	2.32	2.29833	.00570
	2.24	2.32		
	2.24	2.43		
40	2.29	2.55	2.38333	.03495
	2.18	2.55		
	2.18	2.55		

Analysis of Variance

Source	df	MS	F
Total(cor)	53	.01392	-
Time	8	.08586	7.6103
Error	45	.001128	-

Appendix B

R Values

Significant Studentized Ranges for a 5% Level New Multiple Range Test

ν_2 \ g	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.46	3.46	3.46	3.47	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.46	3.46	3.46	3.47	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.45	3.46	3.46	3.47	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.47	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.44	3.45	3.46	3.47	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.44	3.45	3.46	3.47	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.43	3.45	3.46	3.47	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.43	3.44	3.46	3.47	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.43	3.44	3.46	3.46	3.47	3.47	3.47
22	2.93	3.08	3.17	3.24	3.29	3.32	3.35	3.37	3.39	3.42	3.44	3.45	3.46	3.47	3.47	3.47
24	2.92	3.07	3.15	3.22	3.28	3.31	3.34	3.37	3.38	3.41	3.44	3.45	3.46	3.47	3.47	3.47
26	2.91	3.06	3.14	3.21	3.27	3.30	3.34	3.36	3.38	3.41	3.43	3.45	3.46	3.47	3.47	3.47
28	2.90	3.04	3.13	3.20	3.26	3.30	3.33	3.35	3.37	3.40	3.43	3.45	3.46	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.40	3.43	3.44	3.46	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.39	3.42	3.44	3.46	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.37	3.40	3.43	3.45	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.36	3.40	3.42	3.45	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.34	3.38	3.41	3.44	3.47	3.61	3.67

Significant Studentized Ranges for a 1% Level New Multiple Range Test

ν_2 \ g	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.0	9.1	9.2	9.3	9.3	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.3	7.4	7.4	7.5	7.5	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.6	6.6	6.7	6.7	6.8	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.1	6.2	6.2	6.3	6.3	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	5.8	5.9	5.9	6.0	6.0	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.6	5.7	5.7	5.8	5.8	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.5	5.5	5.6	5.7	5.7	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.36	5.42	5.48	5.54	5.55	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.24	5.28	5.34	5.38	5.39	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.13	5.17	5.22	5.24	5.26	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.04	5.08	5.13	5.14	5.15	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	4.96	5.00	5.04	5.06	5.07	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	4.90	4.94	4.97	4.99	5.00	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.84	4.88	4.91	4.93	4.94	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.72	4.75	4.80	4.83	4.86	4.88	4.89	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.76	4.79	4.82	4.84	4.85	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.72	4.76	4.79	4.81	4.82	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.69	4.73	4.76	4.78	4.79	4.79	4.79
22	3.99	4.17	4.28	4.36	4.42	4.48	4.53	4.57	4.60	4.65	4.68	4.71	4.74	4.75	4.75	4.75
24	3.96	4.14	4.24	4.33	4.39	4.44	4.49	4.53	4.57	4.62	4.64	4.67	4.70	4.72	4.74	4.74
26	3.93	4.11	4.21	4.30	4.36	4.41	4.46	4.50	4.53	4.58	4.62	4.65	4.67	4.69	4.73	4.73
28	3.91	4.08	4.18	4.28	4.34	4.39	4.43	4.47	4.51	4.56	4.60	4.62	4.65	4.67	4.72	4.72
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.54	4.58	4.61	4.63	4.65	4.71	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.46	4.51	4.54	4.57	4.59	4.69	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.39	4.44	4.47	4.50	4.53	4.66	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.35	4.38	4.42	4.45	4.48	4.64	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.26	4.31	4.34	4.38	4.41	4.60	4.68

This table is reproduced from David B. Duncan, "Multiple range and multiple *F* tests," *Biometrics*, Volume 11 (1955), p. 4, with the permission of the author of the article and the editor of *Biometrics*.

Appendix C

Q Values

Upper Percentage Points of the Studentized Range, $q_x = \dfrac{\bar{x}_{max} - \bar{x}_{min}}{s_{\bar{x}}}$

Upper 5% Points

Error *df*	p = number of treatment means																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	18.0	27.0	32.8	37.1	40.4	43.1	45.4	47.4	49.1	50.6	52.0	53.2	54.3	55.4	56.3	57.2	58.0	58.8	59.6
2	6.09	8.3	9.8	10.9	11.7	12.4	13.0	13.5	14.0	14.4	14.7	15.1	15.4	15.7	15.9	16.1	16.4	16.6	16.8
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23
5	3.63	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.09	7.17
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.11	6.20	6.27	6.34	6.40	6.47
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.99	6.06	6.14	6.20	6.26	6.33
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.62	5.71	5.80	5.88	5.95	6.03	6.09	6.15	6.21
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	6.00	6.05	6.11
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.72	5.79	5.85	5.92	5.97	6.03
15	3.01	3.67	4.08	4.37	4.60	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.58	5.65	5.72	5.79	5.85	5.90	5.96
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.72	5.79	5.84	5.90
17	2.98	3.63	4.02	4.30	4.52	4.71	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.55	5.61	5.68	5.74	5.79	5.84
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.32	5.39	5.46	5.53	5.59	5.65	5.70	5.75
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.50	5.54	5.59
30	2.89	3.49	3.84	4.10	4.30	4.46	4.60	4.72	4.83	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.48
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74	4.82	4.91	4.98	5.05	5.11	5.16	5.22	5.27	5.31	5.36
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.16	5.20	5.24
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.48	4.56	4.64	4.72	4.78	4.84	4.90	4.95	5.00	5.05	5.09	5.13
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01

Upper Percentage Points of the Studentized Range, $q_x = \dfrac{\bar{x}_{max} - \bar{x}_{min}}{s_{\bar{x}}}$

Upper 1% Points

Error *df*	p = number of treatmen means																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90.0	135	164	186	202	216	227	237	246	253	260	266	272	277	282	286	290	294	298
2	14.0	19.0	22.3	24.7	26.6	28.2	29.5	30.7	31.7	32.6	33.4	34.1	34.8	35.4	36.0	36.5	37.0	37.5	37.9
3	8.26	10.6	12.2	13.3	14.2	15.0	15.6	16.2	16.7	17.1	17.5	17.9	18.2	18.5	18.8	19.1	19.3	19.5	19.8
4	6.51	8.12	9.17	9.96	10.6	11.1	11.5	11.9	12.3	12.6	12.8	13.1	13.3	13.5	13.7	13.9	14.1	14.2	14.4
5	5.70	6.97	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65
8	4.74	5.63	6.20	6.63	6.96	7.24	7.47	7.68	7.87	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.32	8.41	8.49	8.57
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.48	7.60	7.71	7.81	7.91	7.99	8.07	8.15	8.22
11	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95
12	4.32	5.04	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.34	7.42	7.48	7.55
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.12	7.20	7.27	7.33	7.39
15	4.17	4.83	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26
16	4.13	4.78	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.80	6.87	6.94	7.00	7.05
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.96
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.52	6.59	6.65	6.71	6.76	6.82
24	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41
40	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60	5.69	5.77	5.84	5.90	5.96	6.02	6.07	6.12	6.17	6.21
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79	5.84	5.89	5.93	5.98	6.02
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	5.61	5.66	5.71	5.75	5.79	5.83
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65

Source: This table is abridged from Table 29, *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University Press, 1954. It is reproduced with permission of the *Biometrika* trustees and the editors, E. S. Pearson and H. O. Hartley. The original work appeared in a paper by J. M. May, "Extended and corrected tables of the upper percentage points of the 'Studentized' range," *Biometrika*, 39:192–193 (1952).

VITA

Donna Meri Baranowski

Candidate for the Degree of

Master of Science

Report: A Discussion of an Empirical Bayes Multiple Comparison Technique

Major Field: Applied Statistics and Computer Science

Biographical Information:

Personal Data: Date of Birth: November 14, 1947
Height: 5'6" Weight: 120 lbs.
U. S. Citizen
Interests: Tennis, skiing, sailing, personal growth, development, and advancement

Education: M.S., 1978, Utah State University, Logan, Utah
Major: Applied Statistics and Computer Science

Course Work:	Quarter Hours
Programming Fortran	3
Statistical Methods	15
Theory of Probability	11
Analysis Var-Covariance	3
Cluster Analysis	3
Linear Stat Models	3
Intro. Linear Analysis	3
Non-Parametric Statistics	3
Pract. Statistical Consulting	3
Multivariate Analysis	3
Survey Sampling	5
Applied Stochastic Processes	3
Econometrics	3

B.S., 1969, College of Racine, Racine, Wisconsin
Major: Chemistry Minor: Math

Courses taken while with G. D. Searle:

Management of Quality Control and Quality Control Statistics, given by the American Management Association

Environmental Pollution, given by the National Bureau of Standards

Basic Human Interaction, given by the National Training Laboratories

Long Range Planning, by Lewis Allen

Management Development, Kepner Tregoe, Salary Administration, Interviewing, and Hiring, all given by G. D. Searle

Professional Experience: (Employment)

Sept 1978 to present — Corning Glass Works
Senior Quality Control Supervisor

Responsible for the quality and process control in the Electronic Materials Plant of the Technical Products Division. Including the supervision of both hourly and technical personnel.

Mar 1977 to Aug 1978 — Texas A&M University
Lab Technician to B. J. Presley, Chemical Oceanography Department

Responsible for flame and flameless Atomic Absorption analysis and data reduction of trace metals found in fish specimens for the Bureau of Land Management's offshore baseline study of the Gulf of Mexico. Includes on board cruise experiences in the Atlantic and Caribbean waters.

June 1976 to Feb 1977 — Utah State University
Research Assistant to James MacMahon, Ecology Department

Responsible for lab management and nitrogen analysis in connection with the Desert Biome project, a National Science Foundation grant. Also line transect work in the Senoron desert.

1970 to Aug 1975	Searle New Venture Division of G. D. Searle Manager of Quality Assurance Responsible for analytical enzyme and food analysis on all incoming and manufactured goods. Furnished analytical services needed by other departments. Supervised nine chemists and technicians. Accountable for yearly budget and justification of capital expenditures. Established procedures for recording analysis and filing. Adapted published analytical methods for lab use. Held accountable for certificates of analysis on manufactured goods. Interacted with marketing by training technical sales representatives and setting quality control specifications. Lab Instrumentation: Gas and Liquid Chromatography, ultraviolet, visible, and infrared Spectrophotometry, Refractometry, and Polarimetry.
June 1969 to 1970	Hired as a control chemist in the Searle Quality Assurance Lab

References: Will be furnished upon request

Address: 1204 Airline Drive
College Station, Texas 77840
(713) 693-0629 (Home)
(713) 845-5136 (Business)